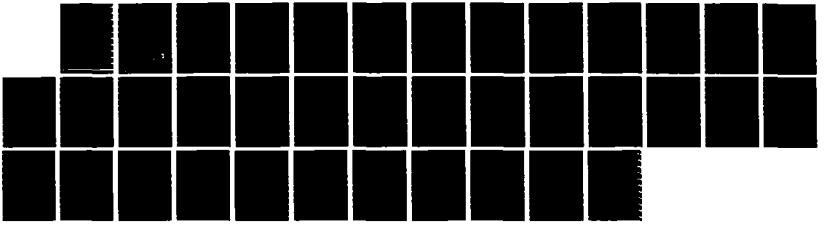
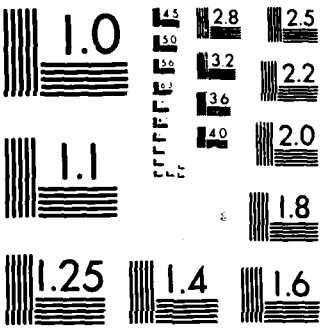
AD-A171 859 IID REPRESENTATIONS FOR THE BIVARIATE PRODUCT LIMIT ESTIMATORS AND THE BO..(U) CALIFORNIA UNIV DAVIS INTERCOLLEGE DIV OF STATISTICS S LO ET AL. SEP 85 1/1

UNCLASSIFIED UCD-STATISTICS-69 AFOSR-TR-86-0689 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A171 859

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS:
2a. SECURITY CLASSIFICATION AUTHORITY: NA	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): UCD STATISTICS REPORT No. 69 AFOSR-85-0268	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR.TR. 86-0689
6a. NAME OF PERFORMING ORGANIZATION: University of California, Davis	6b. OFFICE SYMBOL (If applicable):
6c. ADDRESS (City, State and ZIP Code): Division of Statistics University of California Davis, CA 95616	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM 7b. ADDRESS (City, State and ZIP Code): Bldg.410 Bolling AFB,DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR	8b. OFFICE SYMBOL (If applicable): NM 9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-85-0268
8c. ADDRESS (City, State and ZIP Code): BLDG. 410 Bolling AFB,DC	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A5	

11. TITLE (Include Security Classification): I.I.D. Representations for the Bivariate... (see 16)

12. PERSONAL AUTHOR(S): Shaw-Hwa Lo and Jane-Ling Wang

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical Report	FROM _____ TO _____	September 1985	34

16. SUPPLEMENTARY NOTATION: I.I.D. Representations for the Bivariate Product Limit Estimators and the Bootstrap Versions

17. COSATI CODES		
FIELD	GROUP	SUB. GR.
XXXXXXXXXXXXX		

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number): Bivariate censored data, Product limit estimator, function weak convergence, function LIL and bootstrap.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

see reverse side

DTIC FILE COPY

DTIC ELECTE SEP 15 1986 S D E

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Jane-Ling Wang	(916)752-7624	AFOSR/NM

Let $F(s,t) = P(X>s, Y>t)$ be the bivariate survival function which is subject to random censoring. Let $\hat{F}_n(s,t)$ be the bivariate product limit estimator (PL-estimator) by Campbell and Földes (1982 Proceedings of the International Colloquium on Non-parametric Statistical Inference, Budapest 1980, North Holland). In this paper, it was shown that $\hat{F}_n(s,t) - F(s,t) = n^{-1} \sum_{i=1}^{n} \zeta_i(s,t) + R_n(s,t)$, where $\{\zeta_i(s,t)\}$ is i.i.d. mean zero process and $R_n(s,t)$ is of the order $O((n^{-1}\log n)^{3/4})$ a.s. uniformly on compact sets. Tightness of the process $\{n^{-1} \sum_{i=1}^{n} \zeta_i(s,t)\}$ is shown which implies the weak convergence of the process to a two-dimensional-time Gaussian Process whose covariance structure is given. Corresponding results are also derived for the bootstrap estimators. The results can be extended to the multivariate cases and are extensions of the univariate case of Lo and Singh (1985, Tech. Report, Dept. of Statistics, Rutgers University). Since $\hat{F}_n(s,t)$ may not be a survival function, it is modified to be one. The modified estimator is closer to the true survival function than the bivariate PL-estimator in supnorm.

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

QUALITY INSPECTED 1

AFOSR-TR- 86-0689

I.I.D. Representations for the Bivariate Product Limit Estimators and the bootstrap versions.

by

Shaw-Hwa Lo and Jane-Ling Wang[1]

Rutgers University and University of California, Davis

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

Approved for public release;
distribution unlimited.

AMS (1980) Subject classification: 62G05, 62E20

Keywords and phrases: Bivariate censored data, Product limit estimator, functional weak convergence, function LIL and bootstrap.

[1]Research support in part by Air Force grant AFOSR-85-0268.

86 9 15 055

ABSTRACT

I.I.D. Representations for the Bivariate Product Limit Estimators and the bootstrap versions.

by

Shaw-Hwa Lo
Department of Statistics
Rutgers University

and

Jane-Ling Wang
Division of Statistics
University of Calif. Davis

Let $F(s,t) = P(X > s, Y > t)$ be the bivariate survival function which is subject to random censoring. Let $\hat{F}(s,t)$ be the bivariate product limit estimator (PL- estimator) by Campbell and Földes (1982, Proceedings of the International Colloquium on Non-parametric Statistical Inference, Budapest 1980, North Holland). In this paper, it was shown that $\hat{F}_n(s,t) - F(s,t) = n^{-1} \sum_{i=1}^{n} \zeta_i(s,t) + R_n(s,t)$, where $\{\zeta_i(s,t)\}$ is i.i.d. mean zero process and $R_n(s,t)$ is of the order $O((n^{-1}\log n)^{3/4})$ a.s. uniformly on compact sets. Tightness of the process $\{n^{-1/2} \sum_{i=1}^{n} \zeta_i(s,t)\}$ is shown which implies the weak convergence of the process to a two-dimensional-time Gaussian Process whose covariance structure is given. Corresponding results are also derived for the bootstrap estimators. The results can be extended to the multivariate cases and are extensions of the univariate case of Lo & Singh (1985, Tech. Report, Dept. of Statistics, Rutgers University). Since $\hat{F}_n(s,t)$ may not be a survival function, it is modified to be one. The modified estimator is closer to the true survival function than the bivariate PL-estimator in supnorm.

1. Introduction and Summary.

The one-dimensional product limit estimator (PL-estimator) by Kaplan & Meier (1958) has been treated extensively; e.g. Breslow & Crowley (1974). Csörgő & Horváth (1983), Gill (1983) and etc. Recently Lo & Singh (1985); hereafter abbreviated as [L&S] (1985); represent the PL-process as mean of i.i.d. processes with a remainder of the order $O((n^{-1}\log n)^{3/4})$(a.s.) uniformly on a compact interval. The functional law of iterated logarithm (LIL) and the weak convergence of the PL-estimator by Breslow & Crowley (1974) are all direct applications of this i.i.d. representation.

The bivariate (or more generally, the multivariate) random censoring model did not appear in the literature until near 1980.

Let (X_i^0, Y_i^0), $i = 1,\ldots,n$ be the lifetime vectors of n pairs of items which are i.i.d. with continuous survival function $F(s,t) = P(X^0 > s, Y^0 > t)$. Let (C_i, D_i) be the vector of censoring times of (X_i^0, Y_i^0) and (C_i, D_i) $i = 1,\ldots,n$ are i.i.d. with survival function $G(s,t) = P(C > s, D > t)$. In the bivariate random censorship model, one observed $(X_i, Y_i, \delta_{1i}, \delta_{2i})$, $i = 1,\ldots,n$, where

$X_i = \min\{X_i^0, C_i\}$, $Y_i = \min\{Y_i^0, D_i\}$,

$\delta_{1i} = 1$ or 0 according as $X_i^0 \leq C_i$ or $X_i^0 > C_i$, and

$\delta_{2i} = 1$ or 0 according as $Y_i^0 \leq D_i$ or $Y_i^0 > D_i$.

It is assumed that (X_i^0, Y_i^0) and (C_i, D_i) are mutually independent, for $i = 1,\ldots,n$. Let $H(s,t) = P(X_i > s, Y_i > t)$ denote the survival function of (X_i, Y_i). By independence, $H(s,t) = F(s,t)\,G(s,t)$. Based on the observations $(X_i, Y_i, \delta_{1i}, \delta_{2i})$

one would like to estimate F(s,t).

Situations where estimators of the bivariate survival functions are needed in the presence of random censorship arise commonly in medical follow up for paired observations (eyes, kidneys, matched pair of treatment and control objects, and (pre-test, post-test) situations), and in engineering for two-component system of which the two components are dependent on each other. Another potential application is the stochastic regression model where both the independent and dependent variables are subject to random censoring. Campbell (1981) developed a bivariate self-consistent estimator for discrete times of deaths or losses. A self-consistent approach for the continuous case has been suggested by Korwar & Dahiya (1982). Under the condition $C_i = D_i$, Munoz (1980a, 1980b) shows how to compute the two-dimensional generalization of the PL-estimator through algorithms and proves it is the generalized maximum likelihood estimator and its consistency. Hanley and Parnes (1983) use the EM algorithm to treat maximum likelihood approaches to bivariate estimations. Tsai, Leurgans and Crowley (1983) also construct a family of nonparametric bivariate estimators from a decomposition of the bivariate survival function and shows their consistency. Campbell & Földes (1982); hereafter abbreviate as [C&F] (1982); construct two path dependent bivariate PL-type estimators of the true survival function F(s,t) and establish their rates of strong uniform consistency on a compact set. Noted that the estimators may not be survival functions (see section 6 of [C&F] (1982)). Later on Campbell (1982) shows the weak convergence of these bivariate PL-estimators to a

Gaussian process. Horváth (1983) generalizes the results of [C&F] (1982) to the multivariate case and obtains the rate of convergence on the Euclidean space $\mathbb{R}^k$.

It is the purpose of this paper to further investigate the path dependent bivariate (multivariate) PL-estimator of [C&F](1982) (Horváth (1983)). For simplicity, we shall focus on the bivariate case. The multivariate case can be dealt with similarly. Two path-dependent PL-estimators are introduced in Section 2 of [C&F](1982). We shall consider only one of them in this paper as the other can be treated with symmetric argument. We shall estimate $F(s,t)$ based on the fact that $F(s,t) = F(s,0)F(t|s)$, where $F(t|s) = P(Y^0 > t | X^0 > s)$.

Let
$$N_n(s,t) = N(s,t) = \sum_{i=1}^{n} I(X_i > s,\ Y_i > t)$$

$$\alpha_i(s,t) = I(X_i < s,\ Y_i > t,\ \delta_{1i} = 1),\ i = 1,2,\ldots,n,$$

$$\beta_j(s,t) = I(X_j > s,\ Y_j < t,\ \delta_{2j} = 1),\ j = 1,2,\ldots,n,$$

where $I(\cdot)$ is the indicator function. To estimate $F(s,o)$, project all points (X_i,Y_i) vertically onto the X-axis and ignore the Y_i values. Let $\hat{F}_n(s,0)$ be the one-dimensional PL-estimator of $F(s,0)$ based on (X_i, δ_{1i}), $i = 1,\ldots,n$. That is ,

$$\hat{F}_n(s,0) = \begin{cases} \prod_{i=1}^{n} \left(\dfrac{N(X_i,0)}{N(X_i,0)+1} \right)^{\alpha_j(s,0)} & \text{if } s < X_{(n)} \\ 0 & \text{otherwise,} \end{cases}$$

where $X_{(n)} = \max_{1 \le i \le n} \{X_i\}$.

To estimate $F(t|s)$, project all points (X_i, Y_i) for which $X_i > s$ horizontally to the line $X = s$, and ignore the X_i values. Let $\hat{F}_n(t|s)$ be the one-dimensional PL-estimator of $F(t|s)$ based on (Y_i, δ_{2i}), for which $X_i > s$. That is,

$$\hat{F}_n(t|s) = \begin{cases} \prod_{j=1}^{n} \left(\frac{N(s,Y_j)}{N(s,Y_j)+1} \right)^{\beta_j(s,t)} & \text{if } t < Y_{(n)}(s) \\ 0 & \text{otherwise,} \end{cases}$$

where $Y_{(n)}(s) = \max_{1 \le i \le n} \{Y_i : X_i > s\}$.

Our estimator of $F(s,t)$ is $\hat{F}_n(s,t) = \hat{F}_n(s,0)\hat{F}_n(t|s)$, which is the product of two one-dimensional PL-estimator. Note that $\hat{F}_n(s,t)$ may not be a survival function as mentioned earlier.

Let (S,T) be any fixed point in $\mathbb{R}^2$ such that $H(S,T) > 0$. In this paper, we shall consider the behavior of the bivariate PL-process $n^{1/2}(\hat{F}_n(s,t) - F(s,t))$ on the compact rectangle $[0,S] \times [0,T]$. The results of [L&S] (1985) are extended to the multivariate case. Since the i.i.d. representation of $\hat{F}_n(s,t) - F(s,t)$ in section 3 originates from the corresponding univariate case in [L&S] (1985), the relevant results are summarized in Section 2.

In Theorem 3.1 the bivariate PL-process $\hat{F}_n(s,t) - F(s,t)$ is expressed as $n^{-1} \sum_{i=1}^{n} \zeta_i(s,t)$ plus a remainder of the order $O((n^{-1}\log n)^{3/4})$ (a.s.) uniformly on $[0,S] \times [0,T]$, where $\zeta_i(s,t)$ are i.i.d. random variables. The mean and covariance of the process $\{\zeta(s,t)\}$ are given in Proposition 3.1.

Let $\bar{\zeta}(s,t) = n^{-1} \sum_{i=1}^{n} \zeta_i(s,t)$. Tightness of the process $\{n^{1/2}\bar{\zeta}(s,t)\}$ is

shown on compact sets (Corollary 4.1), and therefore implies the weak convergence of the bivariate PL-process $n^{1/2}(\hat{F}_n(s,t) - F(s,t))$ to a two-dimensional-time Gaussian process. Noted that although the weak convergence of the bivariate PL-process has been shown in Campbell (1982), the covariance structure of the limiting Gaussian process is first given in this paper in Proposition 3.1. The LIL on compact sets also follows directly which generalizes the results of [C&F] [1982]. The proof of tightness in the bivariate case is much harder than the univariate case in [L&S] (1985), and utilizes the tightness results of Bickel & Wichura (1971). In Sections 3 and 4 similar results are also obtained for the bootstrap bivariate PL-process (defined in Section 3), which provide a way to estimate the standard error of $\hat{F}_n(s,t)$ or to give a confidence interval (band) for $F(s,t)$.

Since the bivariate PL-estimator $\hat{F}_n(s,t)$ may not be a survival function, a modified estimator $\tilde{F}_n(s,t)$ which is a survival function itself, is constructed in Section 5. The modified estimator $\tilde{F}_n(s,t)$ is shown in Theorem 5.1 to be closer to the true life distribution than $\hat{F}_n(s,t)$.

All the results in Sections 3 to 5 are generalized to the multivariate case in Section 6.

Some of the lengthly proofs are relegated in the appendices.

While this paper deals only with the multivariate random censoring model, it is possible to extend the results to the multivariate competing risk models.

2. Preliminaries

Since the i.i.d. representation of $\hat{F}_n(s,t) - F(s,t)$ in section 3 originates from the corresponding univariate case in [L&S] (1985), for readers' convenience we shall summarize the relevent results of the univariate case in this section.

Let us first define some notations for the univariate case. Let $\{X_i^0\}$, $i = 1,\ldots,n$ be i.i.d. with continuous survival function $F(t) = P(X_i^0 > t)$, and $\{C_i\}$, $i = 1,\ldots,n$ be i.i.d. with survival function $G(t) = P(C_i > t)$, where C_i and X_i^0 are independent for each i. In the univariate random censorship model one observes (X_i,δ_i), where $X_i = \min(X_i^0,C_i)$ and $\delta_i = 1$ or 0 according as $X_i^0 < C_i$ or $X_i^0 > C_i$. Let $H(t) = P(X_i > t)$ and $H_1(t) = P(X_i^0 > t \text{ and } \delta_i = 1)$. For any positive reals x,t, and δ taking values 0 or 1, define $\phi(x,\delta,t) = -[g(x \wedge t) + H(x)^{-1} I(x < t \text{ and } \delta = 1)]$, where $g(y) = \int_0^y [H(s)]^{-2} d H_1(s)$ and $x \wedge t$ means the minimum of x and t. Let $\hat{F}_n(t)$ be the PL-estimator of F(t), and Δ be any point such that $F(\Delta) > 0$. For $t < \Delta$, it can be checked without difficulties that $\phi(x,\delta,t)$ is a uniform bounded random variable with mean zero and $\mathrm{Cov}(\phi(X,\delta,s), \phi(X,\delta,t)) = -g(s \wedge t)$.

Next, we shall define the bootstrap sample of the PL-estimator $\hat{F}_n(t)$. As mentioned in Efron (1981) we shall adopt the method of drawing random samples (with replacement) (X_i^*, δ_i^*), $i = 1,\ldots,n$, from the population $\{(X_i, \delta_i), i = 1,\ldots,n\}$ giving each doublet equal chance (1/n) at each draw and constructing the PL-estimator $\hat{F}_n^*(t)$ using the bootstrap sample (X_i^*, δ_i^*). The histogram of a large number of values for $n^{1/2}(\hat{F}_n^* - \hat{F}_n)$ is used to approximate the distribution of $n^{1/2}(\hat{F}_n - F)$. Proposition 2.1 provides a first order consistent approximation and hence shows the validity of the bootstrap method.

Proposition 2.1. If F is continuous, one has the following i.i.d representations:

(a) $\log \hat{F}_n(t) - \log F(t) = n^{-1} \sum_{i=1}^{n} \phi(X_i,\delta_i,t) + \gamma_n(t)$, where

$\sup_{0\le t\le\Delta} |\gamma_n(t)| = O((n^{-1}\log n)^{3/4})$ a.s.

(b) $\hat{F}_n(x) - F(x) = n^{-1}F(t) \sum_{i=1}^{n} \phi(X_i,\delta_i,t) + \gamma_n'(t)$,

where $\sup_{0\le t\le\Delta} |\gamma_n'(t)| = O((n^{-1}\log n)^{3/4})$ a.s.

(c) $\hat{F}_n^*(x) - \hat{F}(x) = n^{-1}F(t) \sum_{i=1}^{n} [\phi(X_i^*, \delta_i^*,t) - \phi(X_i,\delta_i,t)] + \gamma_n^*(t)$

where $\sup_{0\le t\le\Delta} |\gamma_n^*(t)| = O((n^{-1}\log n)^{3/4})$ a.s. (p^*), and p^* stands for P in bootstrap probability.

Proof: (b) and (c) are Theorem 1 of [L&S] (1985), and (a) follows from the proof of Theorem 1. □

The i.i.d. representations in Proposition 2.1 ((a) and (b)) provide a way to study the large sample properties of $\hat{F}_n(t)$ and its hazard function $-\log \hat{F}_n(t)$. As a result, weak convergence and law of iterated logarithm (LIL) follow immediately. Proposition (2.1)(c) shows the feasibility of the bootstrap method for estimating the standard error of the PL-estimator $\hat{F}_n$ and constructing confidence band for F(t).

3. I.I.D. Representations.

In this section, the results of Proposition 2.1 are first extended in Lemma 3.1 to conditionnal survival functions and then in Theorem 3.1 to bivariate survival functions. For the bivariate case, we shall consider the same bootstrap method as in Section 2 of drawing random samples (with replacement) $(X_i^*, Y_i^*, \delta_{1i}^*, \delta_{2i}^*)$, $i = 1,\ldots,n$ from the population $\Omega^* = \{(X_i, Y_i, \delta_{1i}, \delta_{2i}); i = 1,\ldots,n\}$, giving each element in Ω^* equal chance $(1/n)$ at each draw. The bivariate PL-estimator $\hat{F}_n^*(s,t)$ is then constructed as $\hat{F}_n(s,t)$ but using the bootstrap sample $\{(X_i^*, Y_i^*, \delta_{1i}^*, \delta_{2i}^*)\}_{i=1,\ldots,n}$ instead; thus $\hat{F}_n^*(s,t) = \hat{F}_n^*(s,0)\ \hat{F}_n^*(t|s)$.

We shall adopt the notations of Section 1 for the bivariate censor model, and define $H(t|s) = P(Y > t|X > s)$, $H_1(t|s) = P(Y > t,\ \delta_2 = 1|\ X > s)$, $H_{1x}(s) = P(X > s,\ \delta_1 = 1)$. For positive reals x,y,s,t, and δ_1,δ_2 taking values 0 or 1, let

$$\xi(x,\delta_1,s) = -\,[g(x\wedge s) + [H(x,0)]^{-1}I(x\leq s \text{ and } \delta_1 = 1)], \text{ where } g(x) = \int_0^x [H(s,0)]^{-2} dH_{1x}(s),$$

and

$$\xi_s(y,\delta_2,t) = -[g_s(y\wedge t) + [H(y|s)]^{-1}I(y\leq t,\ \delta_2 = 1)], \text{ where } g_s(u) = \int_0^u [H(y|s)]^{-2} dH_1(y|s).$$

Let (S,T) be any point with $H(S,T) > 0$.

Let $m_s = \sum_{i=1}^{n} I(X_i > s)$, $m_s^* = \sum_{i=1}^{n} I(X_i^* > s)$.

Lemma 3.1. The following is true if $F(s,t)$ is continuous.

(a) $\log \hat{F}_n(t|s) - \log F(t|s) = m_s^{-1} \sum_{i=1}^{n} \xi_s(Y_i, \delta_{2i}, t) I(X_i > s)$

$+ R_n(t|s)$, where $\sup_{0<s<S,\ 0<t<T} |R_n(t|s)| = O((n^{-1}\log n)^{3/4})$ a.s.

(b) $\log \hat{F}_n^*(t|s) - \log \hat{F}_n(t|s) = m_s^{*-1} \sum_{i=1}^{n} [\xi_s(Y_i^*, \delta_{2i}^*, t) I(X_i^* > s) -$

$\xi_s(Y_i, \delta_{2i}, t) I(X_i > s)] + R_n^*(t|s)$, where

$\sup_{0<s<S,\ 0<t<T} |R_n^*(t|s)| = O((n^{-1}\log n)^{3/4})$ a.s. (P^*), where P^*

stands for the bootstrap probability.

Proof: The proof is similar to that of Theorem 1 of [L&S](1985) and is given in Appendix I. □

Let $\eta(x,y,\delta_1,\delta_2,s,t) = \xi(x,\delta_1,s) + [H(s,0)]^{-1}\xi_s(y,\delta_2,t)I(x > s)$. Define $H^*(t|s)$, $H_1^*(t|s)$, $H_{1x}^*(s)$ similarly by using the bootstrap sample $(X^*, Y^*, \delta_1^*, \delta_2^*)$ instead.

Theorem 3.1 If $F(s,t)$ and $G(s,0)$ are continuous, we have

(a) $\log \hat{F}_n(s,t) - \log F(s,t) = n^{-1} \sum_{i=1}^{n} \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s, t) + R_n(s,t)$ where

$\sup_{0<s<S,\ 0<t<T} |R_n(s,t)| = O((n^{-1}\log n)^{3/4})$ a.s.

(b) $\hat{F}_n(s,t) - F(s,t) = n^{-1}F(s,t) \sum_{i=1}^{n} \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s, t) + R_n'(s,t)$ where

$\sup_{0<s<S,\ 0<t<T} |R_n'(s,t)| = O((n^{-1}\log n)^{3/4})$ a.s.

(c) $\hat{F}_n^*(s,t) - \hat{F}_n(s,t) = n^{-1}F(s,t) \sum_{i=1}^{n} [\eta(X_i^*, Y_i^*, \delta_{1i}^*, \delta_{2i}^*, s,t)$

$- \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s,t)] + R_n^*(s,t)$ where

$$\sup_{0<s<S,\ 0<t<T} |R_n^*(s,t)| = O((n^{-1}\log n)^{3/4}) \text{ a.s. } (p^*).$$

Proof of (a):

Proposition 2.1(a) and Lemma 3.1 imply that

$$\log \hat{F}_n(s,t) - \log F(s,t) = [\log \hat{F}_n(s,0) - \log F(s,0)] + [\log \hat{F}_n(t|s) - \log F(t|s)]$$

$$= n^{-1} \sum_{i=1}^{n} \xi(X_i, \delta_{1i}, s) + m_s^{-1} \sum_{i=1}^{n} \xi_s(Y_i, \delta_{2i}, t)I(X_i > s) + R_{n1}(s,t), \text{ where}$$

$$\sup_{0<s<S,\ 0<t<T} |R_{n1}(s,t)| = O((n^{-1}\log n)^{3/4}) \text{ a.s.}$$

$$= n^{-1} \sum_{i=1}^{n} \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s,t) + R_{n1}(s,t) + R_{n2}(s,t), \text{ where}$$

$$R_{n2}(s,t) = \{n/m_s - [H(s,0)]^{-1}\}\ n^{-1} \sum_{i=1}^{n} \xi_s(Y_i, \delta_{2i}, t)I(X_i > s).$$

It is easy to see at this stage that, $R_{n2}(s,t) = O(n^{-1}\log \log n)$ a.s. for each (s,t). To show that it holds uniformly for $0<s<S$, $0<t<T$, we shall apply the functional LIL due to Dudley & Philipp (1983 Theorem 4.1).

Let $Z_i = \xi_s(Y_i, \delta_{2i}, t)I(X_i > s)$, Z_i takes values in $D[0,S] \times D[0,T]$ under the sup norm $\| \ \|$ on $[0,S] \times [0,T]$, and $S_n = \sum_{j=1}^{n} Z_j$. It is clear that $E\|Z_1\|^2 < \infty$, and

hence Condition (4.2) of Dudley & Philipp is satisfied.

Condition (4.1) is satisfied due to the tightness of the process $n^{1/2}S_n$ which is shown in Theorem 4.1 of the next section. It then follows from Theorem 4.1 of Dudley & Philipp (1983) that $\|S_n/n\| = O((n^{-1}\log\log n)^{1/2})$ a.s.

Also $\sup_{0\le s\le S} |n/m_s - [H(s,0)]^{-1}| = 0\ ((n^{-1}\log\log n)^{1/2})$ a.s. from the LIL for empirical distribution and the fact that $H(S,0) > 0$.

We have thus shown that

$$\sup_{0\le s\le S,\ 0\le t\le T} |R_{n2}(s,t)| = O(n^{-1}\log\log n) \text{ a.s. .}$$

(a) follows with $R_n(x,t) = R_{n1}(s,t) + R_{n2}(s,t)$.

Proof of (b) : Let $Z_i = \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s,t)$.

It can be checked easily that Z_i is uniformly bounded on $[0,S] \times [0,T]$. Applying Theorem 4.1 of Dudley & Philipp (1983) once more, we have

$$\sup_{0\le s\le S,\ 0\le t\le T} |n^{-1} \sum_{i=1}^{n} \eta(X_i, Y_i, \delta_{1i},\delta_{2i}, s,t)| = O((n^{-1}\log\log n)^{1/2}) \text{ a.s.}$$

(b) then follows from (a) and the two-term Taylor's expansion of $\hat{F}_n(s,t) - F(s,t) = \exp[\log \hat{F}_n(s,t)] - \exp[\log F(s,t)]$.

Proof of (c): Using Proposition 2.1(c) and Lemma 3.1(b) the proof follows by mimicing the proof of (a) and (b). □

The following LIL follows from the proof of Theorem 3.1 by applying Theorem 4.1 of Dudley & Philipp (1983).

Corollary 3.1 If $H(s,t)$ is continuous, we have

(a) $$\sup_{0\le s\le S,\ 0\le t\le T} |\hat{F}_n(s,t) - F(s,t)| = O((n^{-1}\log\log n)^{1/2}) \text{ a.s.}$$

(b) $$\sup_{0\le s\le S,\ 0\le t\le T} |\hat{\hat{F}}_n(s,t) - \hat{F}_n(s,t)| = O_p^*((n^{-1}\log\log n)^{1/2}) \text{ a.s.}$$

Noted that Corollary 3.1 (a) was also obtained in [C&F] (1982) but (b) has never been shown to the best of our knowledge.

Let $\eta(s,t) = \eta(X,Y,\delta_1,\delta_2,s,t)$ and $\Gamma((s,t),(s',t')) = \mathrm{Cov}(\eta(s,t),\eta(s',t'))$. The mean and covariance structure of the process $\{\eta(s,t)\}$ is given in the next proposition.

Proposition 3.1 (a) $E(\eta(s,t)) = 0$

(b) Assume $s \le s'$,

$$\Gamma((s,t),(s',t'))$$

$$= -g(s) + [H(s,0)]^{-1} E[\xi_s(Y,\delta_2,t)\xi(X,\delta_1,s')\cdot I(X > s)]$$

$$+ [H(s,0)]^{-1}\{\int_0^t H(y|s')[H(y|s)]^{-2} g_{s'}(y\ \ t')dH_1(y|s)$$

$$- \int_0^t [H(y|s)]^{-1} g_{s'}(y\ \ t')dH_1(y|s') - \int_0^t\ {}^{t'}\ [H(y|s)\ H(y|s')]^{-1}dH_1(y|s')\}.$$

Proof of (a):

$$E(\eta(s,t)) = E[\xi(X, \delta_1, s) + [H(s,0)]^{-1}\xi_s(Y, \delta_2, t)I(X > s)]$$

$$= E(\xi(X, \delta_1, s)) + [H(s,0)]^{-1}E_{Y|X>s}[\xi_s(Y, \delta_2, t)]\cdot P(X > s)$$

$$= 0, \text{ since both expectations above are zero by [L\&S](1985).}$$

Proof of (b): Due to the conditional argument in $\xi_s(Y,\delta_2,t)$, the covariance structure of the bivariate case is much more complicated than the univariate case and is calculated in Appendix II. □

When $s = s'$, the covariance structure in (b) is much simpler, and we have:

Corollary 3.2. (a) $\Gamma((s,t), (s,t')) = -g(s) - g_s(t \wedge t')$

(b) $\text{Var}(\eta(s,t)) = -[g(s) + g_s(t)]$.

Proof: It can be checked easily from the proof in Appendix II that the second term of $\Gamma((s,t), (s',t'))$ vanishes when $s = s'$. The rest of the proof follows immediately. □

4. Weak Convergence.

Let $\eta_i(s,t) = \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s,t)$, $\bar{\eta}(s,t) = n^{-1} \sum_{i=1}^{n} \eta_i(s,t)$, $\zeta_i(s,t) = F(s,t)\eta_i(s,t)$, $\bar{\zeta}(s,t) = F(s,t)\bar{\eta}(s,t)$ and $\eta_i^*(s,t)$, $\bar{\eta}^*(s,t)$, $\zeta_i^*(s,t)$, $\bar{\zeta}^*(s,t)$ be their bootstrap counterparts. We shall prove the tightness of the processes $n^{1/2}\bar{\eta}(s,t)$, $n^{1/2}\bar{\zeta}(s,t)$ and their bootstrap counterparts. Weak convergence of the corresponding processes to Gaussian Processes then follow from the Cramér-Wold device. For any block $E = (s,s'] \times (t,t']$ define

$$X(E) = \eta(s',t') - \eta(s,t') - \eta(s',t) + \eta(s,t)$$

$$X_i(E) = \eta_i(s',t') - \eta_i(s,t') - \eta_i(s',t) + \eta_i(s,t).$$

$$X_n(E) = n^{1/2}[\bar{\eta}(s',t') - \bar{\eta}(s,t') - \bar{\eta}(s',t) + \bar{\eta}(s,t)] = n^{-1/2} \sum_{i=1}^{n} X_i(E)$$

Lemma 4.1. There exists a positive finite measure W on $\mathbb{R}^+ \times \mathbb{R}^+$ such that $E(|X(E)|^2) \leq W(E)$, for any block E in $[0,S] \times [0,T]$.

Proof: The proof is given in Appendix III. □

Let $B = (s_1,s_2] \times (t_1,t_2]$, $C = (s_2,s_3] \times (t_1,t_2]$ by any two neighboring blocks in $[0,S] \times [0,T]$.

Lemma 4.2. $E(|X_n(B)|^2|X_n(C)|^2) \leq 2W(B)W(C) + O(n^{-1})$, where $O(n^{-1})$ is independent of B and C.

Proof: $E(|X_n(B)|^2|X_n(C)|^2)$

$$= n^{-2}E[(\Sigma X_i(B))^2 (\Sigma X_i(C))^2]$$

$$= n^{-2}E([\sum_{i=1}^{n} X_i^2(B) + \sum_{i\neq j} X_i(B)X_j(B)][\sum_{i=1}^{n} X_i^2(C) + \sum_{i\neq j} X_i(C)X_j(C)])$$

$= n^{-2}\{nE(X^2(B)\ X^2(C)) + n(n-1)\ E(X^2(B))\cdot E(X^2(C)) +$

$n(n-1)[E(X(B)X(C))]^2\}$

$\leq 2E(X^2(B))\cdot E(X^2(C)) + O(n^{-1})$, where $O(n^{-1})$ is independent of B and C since X(E) is uniformly bounded for any block E,

$\leq 2W(B)\cdot W(C) + O(n^{-1})$. □

We shall use the convergence criteria for multiparameter stochastic processes by Bickel & Wichura (1971); hereafter abbreviated as [B&W] (1971) to prove the following tightness theorem.

Theorem 4.1. *If* $G(s,0)$ *and* $F(s,t)$ *are both continuous, then the processes* $\{n^{1/2}\bar{\eta}(s,t)\}$ *and* $\{n^{1/2}[\bar{\eta}^*(s,t) - \bar{\eta}(s,t)]\}$ *are both tight on* $[0,S] \times [0,T]$.

Proof: From Appendix III, $W(x,y) = \text{constant}\cdot H(x,0)H_1(0,y) + \text{constant}\cdot H_1(x,y)$, where $H_1(x,y) = P(X>s,\ Y>t,\ \delta_2 = 1)$. The measure W has continuous marginal since F(s,t) and H(s,0) are both continuous under our assumptions. We shall apply Theorem 3 of [B&W](1971).

Replace W(E) by $W_n(E)$, where $W_n(E) = 2^{1/2}W(E) + O(n^{-1/2})$ *and* $O(n^{-1/2})$ *is* independent of the block E in $[0,S] \times [0,T]$. Hence $(X_n, W_n) \varepsilon\, C\ (2,4)$, where $C\ (\beta,v)$ is defined on page 1658 of [B&W](1971).
Theorem 3 and the discussion after it then implies that $X_n(s,t) = n^{1/2}\bar{\eta}(s,t)$is tight. Noted that although $X_n(s,t)$ does not vanish along the lower boundary of $[0,S] \times [0,T]$, it can be modified by subtracting the boundary value of $X_n(s,t)$ from it.

To show the tightness of $n^{1/2}(\bar{\eta}^*(s,t) - \bar{\eta}(s,t))$, recall that Theorem 3.1(c) implies that bootstrapping $\hat{F}_n(s,t)$ amounts to drawing random sample $\{\eta_i^*(s,t),\ i = 1,\ldots,n\}$ from the population which assigns equal mass $(1/n)$ to each of the n functions $\{\eta_i(s,t),\ i = 1,\ldots,n\}$. For any block E let $X^*(E)$, $X_n^*(E)$ be defined as in $X(E)$, $X_n(E)$ by replacing η by $\eta^* - \eta$. Following the proof of Lemma 4.1 one can show that $E(|X^*(E)|^2) \leq W(E) + O((n^{-1}\log n)^{1/2})$ where $O((n^{-1}\log n)^{1/2})$ is independent of the block E. Lemma 4.2 then implies that $E(|X_n^*(B)|^2 |X_n^*(C)|^2) \leq 2W(B)W(C) + O((n^{-1}\log n)^{1/2})$ a.s. The rest of the proof is similar to that of $n^{1/2}\bar{\eta}(s,t)$. □

The following corollary is immediate.

Corollary 4.1. Under the assumption that $G(s,0)$ and $F(s,t)$ are continuous, the process $\{n^{1/2}\bar{\zeta}(s,t)\}$ and $\{n^{1/2}(\bar{\zeta}^*(s,t) - \bar{\zeta}(s,t))\}$ are both tight on $[0,S] \times [0,T]$.

Theorem 4.1 and its Corollary also show that both the bivariate hazard process and PL-process converge weakly to a two parameter mean zero Gaussian process with covariance $\Gamma((s,t), (s',t'))$ and $F(s,t)F(s'.t')\Gamma((s,t),(s',t'))$respectively. To show that $n^{1/2}(\hat{F}_n^* - \hat{F}_n)$ converges to the same Gaussian process $n^{1/2}(\hat{F}_n - F)$, we only have to show that the finite dimensional distribution $n^{1/2}[\bar{\zeta}^*(s_i,t_i) - \bar{\zeta}(s,t)]$, for some $\{(s_i,t_i),\ 1\leq i\leq k\}$, converge to the k-variate Guassian distribution with mean zero and dispersion matrix $\{F(S_i,t_i)F(S_j,t_j)\Gamma((S_i,t_i),(S_j,t_j))\}$. This follows directly from the bootstrap central limit theorem for sample means (Bickel and Freedman (1981) or Singh (1981)). Thus we have proved the following corollary.

Corollary 4.2. Under the assumption that $G(s,0)$ and $F(s,t)$ are continuous. $n^{1/2}(\hat{F}_n(s,t) - F(s,t))$ and $n^{1/2}(\hat{F}_n^*(s,t) - \hat{F}_n(s,t))$ both converge to the two-parameter Gaussian process with mean zero and covariance $F(s,t)F(s',t')\Gamma((s,t),(s',t'))$.

We have thus shown that the bootstrap method works under the bivariate random censorship model, which provides a way to estimate the standard error of $\hat{F}_n(s,t)$ or to construct a confidence interval (band) for $F(s,t)$. This is most valuable since the covariance structure of $\hat{F}_n(s,t)$ is very complicated as shown in Proposition 3.1.

5. Modified estimators.

The bivariate PL-estimator $\hat{F}_n(s,t)$ defined in section 1 may not be a survival function. Examples are given in [C&F] (1982) and Campbell (1981). What happen is that although $\hat{F}_n(s,t)$ is a non-increasing function of t when s is kept fixed, it may not be nonincreasing function of s when t is kept fixed. We shall modify $\hat{F}_n(s,t)$ to be a survival function and this modified estimator is closer to the true survival function $F(s,t)$ then $\hat{F}_n(s,t)$ in Supnorm distance.

For any $0 \le s < \infty$, $0 \le t < \infty$, define $\tilde{F}_n(s,t) = \sup_{s' \ge s} \hat{F}_n(s',t)$. Note that $\tilde{F}_n(s,t)$ is a step function and $\tilde{F}_n(s,t) = \hat{F}_n(s,y_i) = \sup_{x_k \ge s} \hat{F}_n(x_k,y_i)$, for $y_i \le t < y_{i+1}$.

Proposition 5.1. $\tilde{F}_n(s,t)$ is a survival function.

Proof: For $s' \ge s$, it is clear that $\tilde{F}_n(s,t) \ge \tilde{F}_n(s',t)$.

For $t' \ge t$, $\tilde{F}_n(s,t) = \sup_{s' \ge s} \hat{F}_n(s',t) \ge \sup_{s' \ge s} \hat{F}_n(s',t') = \hat{F}_n(s',t')$. □

For each s, let x_s be the x_j such that $\tilde{F}_n(s,t) = \hat{F}_n(x_j,t) = \sup_{x_k \ge s} \hat{F}_n(x_k,t)$.

Since $\hat{F}_n(s,t) \leq \tilde{F}_n(s,t)$ and $F(s,t) \geq F(x_s,t)$, we obtain the following:

$$\hat{F}_n(s,t) - F(s,t) \leq \tilde{F}_n(s,t) - F(s,t) \leq \hat{F}_n(x_s,t) - F(x_s,t).$$

It then follows immediately that

Proposition 5.2. $\sup_{0\leq s,t<\infty} |\tilde{F}_n(s,t) - F(s,t)| \leq \sup_{0\leq s,t<\infty} |\hat{F}_n(s,t) - F(s,t)|$.

Proposition 5.2 implies that, for any $\lambda_n = o(n^{1/2})$,

$$\lambda_n \sup_{s,t} |\tilde{F}_n(s,t) - \hat{F}_n(s,t)| \to 0 \text{ a.s.} \tag{5.1}$$

It is not clear yet whether λ_n can be taken at the rate of $n^{1/2}$. We conjecture that (5.1) hold for $\lambda_n = n^{1/2}$ under some strict uniform monotonicity of $F(s,t)$, and hence $n^{1/2}(\tilde{F}_n(s,t)-F(s,t))$ converges weakly to the same Gaussian process as $n^{1/2}(\hat{F}_n(s,t) - F(s,t))$ does.

6. Generalizations to multivariate survival function.

For multivariate survival function $F(s_1,s_2,\ldots,s_k)$, we can define the multivariate PL-estimator by conditioning argument similar to that of $\hat{F}_n(s,t)$. See Horväth (1983) page 203 for such extension. Let $\hat{F}_n(s_1,\ldots,s_k)$ denote the multivariate PL-estimator so defined. All the results in this paper can be generalized to $\hat{F}_n(s_1,\ldots,s_k)$ with similar arguments.

7. Appendix I: Proof of Lemma 3.1.

Define the following empirical survival functions:

$$H_n(s,t) = n^{-1} \sum_{i=1}^{n} I(X_i > s,\ Y_i > t), \text{ hence } m_s = nH_n(s,0),$$

$$H_n(t|s) = m_s^{-1} \sum_{i=1}^{n} I(X_i > s,\ Y_i > t),\ H_{1n}(t|s) = m_s^{-1} \sum_{i=1}^{n} I(X_i > s,\ Y_i > t, \delta_{2i} = 1).$$

Let $R_{n1}(t|s) = \log \hat{F}_n(t|s) - \int_0^t [H_n(y|s)]^{-1} dH_{1n}(y|s)$,

$$R_{n2}(t|s) = \int_0^t \{[H_n(y|s)]^{-1} - [H(y|s)]^{-1}\} dH_1(y|s) -$$

$$\int_0^t [H(y|s) - H_n(y|s)][H(y|s)]^{-2}\, dH_1(y|s),$$

$$R_{n3}(t|s) = \int_0^t \{[H_n(y|s)]^{-1} - [H(y|s)]^{-1}\}\, d(H_{1n}(y|s) - H_1(y|s)).$$

[Part 1]

Proof of Lemma 3.1(a):

$$H_1(t|s) = P(Y > t,\ \delta_2 = 1 | X > s)$$

$$= P(Y^0 > t,\ Y^0 \leq D |\ X^0 > s,\ C > s)$$

$$= -\int_t^\infty G(y|s) dF(y|s), \qquad \text{where } G(t|s) = P(D > t |\ C > s).$$

Hence $dH_1(t|s) = G(t|s) \cdot dF(t|s)$.

As a result $\log F(t|s) = \int_0^t [F(y|s)]^{-1} dF(y|s)$

$$= \int_0^t [H(y|s)]^{-1} dH_1(y|s).$$

It can be checked easily as in the proof of Theorem 1 of [L&S] (1985) that

$$n^{-1}\Sigma\, \eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s,t) = \int_0^t [H(y|s)]^{-2}[H(y|s) - H_n(y|s)]dH_1(y|s)$$

$$+ \int_0^t [H(y|s)]^{-1} d(H_{1n}(y|s) - H_1(y|s)).$$

Hence $\log \hat{F}_n(t|s) - \log F(t|s) = R_{n1}(t|s) + n^{-1}\,\Sigma\eta(X_i, Y_i, \delta_{1i}, \delta_{2i}, s,t)$

$$+ R_{n2}(t|s) + R_{n3}(t|s).$$

From (8) of [L&S] (1985), for any $0<s<S$, $\sup_{0<t<T} |R_{n1}(t|s)| = O(m_s^{-1})$. a.s.

Since $m_s > m_S$ for any $0<s<S$, and $m_S = O(n)$ a.s., we have

$$\sup_{0<s<S,\ 0<t<T} |R_{n1}(t|s)| = O(n^{-1}) \text{ a.s.} \quad .$$

Next, consider $R_{n2}(t|s)$.

$$|R_{n2}(t|s)| = \left|\int_0^t [H(y|s) - H_n(y|s)]^2 [H_n(y|s)H^2(y|s)]^{-1} dH_1(y|s)\right|$$

$$< \sup_{0<y<T} |H(y|s) - H_n(y|s)|^2 \cdot \inf_{0<y<T} [H_n(y|s)\cdot H^2(y|s)] \qquad (7.1)$$

By triangular inequality,

$$\sup_{\substack{0<s<S,\\ 0<y<T}} |H(y|s) - H_n(y|s)| < \sup_{\substack{0<s<S,\\ 0<y<T}} [H(s,0)]^{-1}[|H(s,y) - H_n(s,y)| + |H_n(s,0) - H(s,0)|]$$

$$= O((n^{-1}\log\log n)^{1/2}) \text{ a.s., from the}$$

law of iterated logarithm for empirical distribution functions

(See Kiefer (1961)).

Since both $H(y|s)$ and $H_n(y|s)$ are bounded away from zero (a.s.) (7.1)

implies that $\sup_{0<s<S\ \ 0<t<T} |R_{n2}(t|s)| = O(n^{-1}\log\log n)$ a.s.

It only remains to prove

$$\sup_{0\le s\le S,\ 0\le t\le T} |R_{n3}(t|s)| = O((n^{-1}\log n)^{3/4}) \text{ a.s.}, \tag{7.2}$$

and this can be done by similar argument of Lemma 2 of [L&S](1985). We shall provide the proof when G is also assumed to be continuous (and hence H is continuous). The argument can be extended to the case when G is arbitrary as the remark after Lemma 2 of [L&S](1985).

Let $k_n = O((n/\log n)^{1/2})$, $y_1 = 0$, $y_{k_{n+1}} = T$. Partition the interval $[0,T]$ into subintervals $[y_i, y_{i+1}]$ $i = 1,\ldots,k_n$, such that $H(0,y_i) - H(0,y_{i+1}) = O((n^{-1}\log n)^{1/2})$, and hence both $H(s,y_i) - H(s,y_{i+1})$ and $H(y_i|s) - H(y_{i+1}|s)$ are $O((n^{-1}\log n)^{1/2})$ uniformly in s. From now on, all the $O(\cdot)$ terms hold uniformly for $0\le s\le S$, $0\le t\le T$.

$$|R_{n3}(t|s)|$$

$$\le \sum_{i=0}^{k_n} \left|\int_{y_i}^{y_{i+1}} \{[H_n(y|s)]^{-1} - [H(y|s)]^{-1} - [H_n(y_i|s)]^{-1} + [H(y_i|s)]^{-1}\} d(H_{1n}(y|s) - H_1(y|s))\right|$$

$$+ \sum_{i=1}^{k_n} \left|\int_{y_i}^{y_{i+1}} \{[H_n(y_i|s)]^{-1} - [H(y_i|s)]^{-1}\}\, d(H_{1n}(y|s) - H_1(y|s))\right|,$$

$$\le 2 \max_{0\le i\le k_n} \sup_{y_i\le y\le y_{i+1}} |[H_n(y|s)]^{-1} - [H(y|s)]^{-1} - [H_n(y_i|s)]^{-1} + [H(y_i|s)]^{-1}|$$

$$+ k_n \sup_{0\le t\le T} |[H_n(t|s)]^{-1} - [H(t|s)]^{-1}| \sup_{1\le i\le k_n} |H_{1n}(y_{i+1}|s) - H_1(y_{i+1}|s) - H_n(y_i|s) + H_1(y_i|s)|,$$

$$\le A + B \quad \text{(say)}$$

To evaluate A, partition $[y_i, y_{i+1}]$ into subintervals $[y_{ij}, y_{i(j+1)}]$, $j = 1,\ldots,a_n$, where $a_n = O((n/\log n)^{1/4})$, $y_{i1} = y_i$, $y_{i\,(a_n+1)} = y_{i+1}$, such that $H(0,y_{ij}) - H(0,y_{i(j+1)}) = O((n^{-1}\log n)^{3/4})$, and hence both $H(s,y_{ij}) - H(s,y_{i(j+1)})$ and $H(y_{ij}|s) - H(y_{i(j+1)}|s)$ are $O((n^{-1}\log n)^{3/4})$.

Consider

$$\sup_{y_i<y<y_{i+1}} |[H_n(y|s)]^{-1} - [H(y|s)]^{-1} - [H_n(y_i|s)]^{-1} + [H(y_i|s)]^{-1}|$$

$$\leq \max_{1\leq j\leq a_n} |[H_n(y_{ij}|s)]^{-1} - [H_n(y_i|s)]^{-1} - [H(y_{ij}|s)]^{-1} + [H(y_i|s)]^{-1}|$$

$+ O((n^{-1}\log n)^{3/4})$, by monotonicity of $H_n(y|s)$.

$$= \max_{1\leq j\leq a_n} |[H_n(y_{ij}|s) - H(y_{ij}|s)][H(y_{ij}|s)]^{-2} - [H_n(y_i|s) - H(y_i|s)][H(y_i|s)]^{-2}|$$

$+ O((n^{-1}\log\log n) + O((n^{-1}\log n)^{3/4})$ a.s., by LIL for empirical distribution function and strong law of large number.

$$\leq \max_{1\leq j\leq a_n} [H(S,T)]^{-4}|[H_n(y_{ij}|s) - H(y_{ij}|s)]H^2(s,y_i) - [H_n(y_i|s) - H(y_i|s)]H^2(s,y_{ij})|$$

$$+ O((n^{-1}\log n)^{3/4}) \text{ a.s.}$$

$$\leq \max_{1\leq j\leq a_n} [H(S,T)]^{-4}\{|H_n(y_{ij}|s) - H(y_{ij}|s) - H_n(y_i|s) + H(y_i|s)| + 2|H(s,y_i) - H(s,y_{ij})|$$

$\cdot |H_n(y_i|s) - H(y_i|s)|\} + O((n^{-1}\log n)^{3/4})$ a.s.

$= \max_{1\le j\le a_n} [H(S,T)]^{-4}\{|H_n(s,y_{ij}) - H_n(s,y_i) - H(s,y_{ij}) + H(s,y_i)|/H(s,0)\} + O(n^{-1}\log n)$

$+ O((n^{-1}\log n)^{3/4})$ a.s., by LIL for empirical distribution function. Use the exponential inequality of Lemma 1 in [L&S](1985) to get the following probability bound as in Lemma 2 of [L&S](1985):

$\max_{0\le i\le k_n} \max_{1\le j\le a_n} P(|H_n(s,y_{ij}) - H_n(s,y_i) - H(s,y_{ij}) + H(s,y_i))| > \text{constant} \cdot n^{-3/4}\log n^{3/4})$

$= O(n^{-3})$. It then follows from Bonferroni inequality and the Borel Cantelli Lemma that $A = 0((n^{-1}\log n)^{3/4})$ a.s.

To estimate B, notice that $|H_1(s,y_{i+1}) - H_1(s,y_i)| \le |H(s,y_{i+1}) - H(s,y_i)|$. The rest of the proof is similar to that of A and in fact easier, since one need not to partition the interval $[y_i, y_{i+1}]$ further as in B. We thus have $B = O((n^{-1}\log n)^{3/4})$ a.s. .

(7.2) now follows immediately. □

[Part 2]

Proof of Lemma 3.1(b):

Define $H_n^*(s,t)$, $H_n^*(t|s)$, and $H_{1n}^*(t|s)$ as $H_n(s,t)$, $H_n(t|s)$, $H_{1n}(t|s)$ by using the bootstrap sample instead. For example, $H_{1n}^*(t|s) = [m_s^*]^{-1} \sum_{i=1}^{n} I(X_i^* > s, Y_i^* > t, \delta_{2i}^* = 1)$. By the same argument as the proof of Lemma 3.1(a) above and Theorem 1 in [L&S] (1985), we have the following bootstrap version

$$\sup_{0\le s\le S,\ 0\le t\le T} |\log \hat{F}_n^*(t|s) - \int_0^t [H_n^*(y|s)]^{-1} dH_{1n}^*(y|s)| = O(n^{-1}) \text{ a.s. } (p^*).$$

Since $\sup_{0\le s\le S,\ 0\le t\le T} |R_{n1}(s,t)| = O(n^{-1})$ a.s., we arrive at the following:

$$\log \hat{F}_n^*(t|s) - \log F_n(t|s) = \int_0^t \{[H_n^*(y|s)]^{-1} - [H_n(y|s)]^{-1}\} dH_{1n}^*(y|s)$$

$$+ \int_0^t [H_n(y|s)]^{-1} d(H_{1n}^*(y|s) - H_{1n}(y|s)) + O(n^{-1}) \text{ a.s. } (p^*)$$

$$= I + II + O(n^{-1}) \text{ a.s. } (p^*), \text{ where } O(n^{-1}) \text{ holds uniformly on } [0,S]\times[0,T].$$

Mimicing the proof of (a) for $R_{n3}(t|s)$, we have

$$I = \int_0^t \{[H_n^*(y|s)]^{-1} - [H_n(y|s)]^{-1}\} dH_{1n}(y|s) + O((n^{-1}\log n)^{3/4}) \text{ a.s. } (p^*)$$

$$= \int_0^t [H_n(y|s) - H_n^*(y|s)]\ [H_n(y|s)]^{-2} dH_{1n}(y|s) + O(n^{-1} \log n)$$

$$+ O((n^{-1} \log n)^{3/4}) \text{ a.s. } (p^*)$$

$$= \int_0^t [H_n(y|s) - H_n^*(y|s)][H(y|s)]^{-2} dH_{1n}(y|s) + O((n^{-1}\log n)^{3/4}) \text{ a.s. } (p^*) \quad (7.3)$$

The last two equalities follow by similar argument as $R_{n2}(t|s)$ in Lemma 3.1(a). Mimicing the proof of Lemma 3.1(a) for $R_{n3}(t|s)$ once more, we have

$$II = \int_0^t [H(y|s)]^{-1} d(H_{1n}^*(y|s) - H_{1n}(y|s)) + O((n^{-1}\log n)^{3/4}) \text{ a.s. } (p^*) \quad (7.4)$$

Lemma 3.1.(b) now follows from (7.3) and (7.4) since all the $O(\cdot)$ terms hold uniformly on $[0,S]\times[0,T]$. □

8. Appendix II: Proof of Proposition 3.1(b)

$\mathrm{Cov}(n(s,t),\ n(s',t'))$

$= E(n(s,t)n(s',t'))$, from proposition 3.1(a)

$= E\{[\zeta(X,\delta_1,s) + [H(s,0)]^{-1}\zeta_s(Y,\ \delta_2,t)I(X > s)]\cdot[\zeta(X,\delta_1,s') +$

$[H(s',0)]^{-1}\zeta_{s'}(Y,\ \delta_2,t')I(X > s')]\}$

$= E[\zeta(X,\delta_1,s)\cdot\zeta(X,\delta_1,s')] + [H(s,0)]^{-1}E[\zeta_s(Y,\delta_2,t)I(X > s)\cdot\zeta(X,\delta_1,s')]$

$+ [H(s',0)]^{-1}E[\zeta(X,\ \delta_1,s)\zeta_{s'}(Y,\ \delta_2,t')I(X > s')] + [H(s,0)H(s',0)]^{-1}$

$\cdot E[\zeta_s(Y,\delta_2,t)\cdot\zeta_{s'}(Y,\delta_2,t')I(X > s' \vee s)]$

$= I + II + III + IV$, where $s' \vee s = \max\{s',\ s\}$.

It follows from direct calculation or [L&S] (1985) that $I = -g(s \wedge s') = -g(s)$, for $s < s'$.

Next we shall show that III vanishes for $s < s'$.

To see this, consider

$E[\zeta(X,\delta_1,s)\cdot\zeta_{s'}(Y,\delta_2,t')I(X > s')]$

$= E\{[g(X\wedge s) + [H(X,0)]^{-1}I(X < s,\ \delta_1 = 1)][g_{s'}(Y\wedge t') + [H(y|s')]^{-1}I(y < t',\delta_2 = 1)]$

$\cdot I(X > s')\}$

$= \{E_{Y|X>s'}[g(s)g_{s'}(Y\wedge t')] + E_{Y,\delta_2|X>s'}[g(s)[H(Y|s')]^{-1}I(Y<t',\ \delta_2 = 1)]\}\cdot H(s',0)$

$= \{\int_0^t H(y|s')dg_{s'}(y) - \int_0^{t'} [H(y|s')]^{-1}dH_1(y|s')\}g(s)H(s',0)$

$= 0$, since $dg_{s'}(y) = [H(y|s')]^{-2}dH_1(y|s')$.

Hence III = 0, for $s < s'$.

Since II can not be simplied, we shall leave it as it is. Noted that if $s>s'$, by symmetric argument, II = 0.

It remains to evaluate IV.

$$IV = [H(s,0)H(s',0)]^{-1}E[\zeta_s(Y,\delta_2,t)\zeta_{s'}(Y,\delta_2,t')I(X > s' \vee s)]$$

$$= [H(s,0)H(s',0)]^{-1}H(s',0)E_{Y|X>s'}\{[g_s(Y\wedge t) + [H(y|s)]^{-1}I(Y<t,\ \delta_2=1)]\cdot$$

$$[g_{s'}(Y\wedge t') + [H(Y|s')]^{-1}I(Y<t',\ \delta_2=1)]\}$$

$$= H(s,0)]^{-1}\{E_{Y|X>s'}[g_s(Y\wedge t)g_{s'}(Y\wedge t')] + E_{Y|X>s'}[g_s(Y\wedge t)[H(Y|s')]^{-1}I(Y<t',\delta_2=1)]$$

$$+ E_{Y|X>s'}[g_{s'}(Y\wedge t')[H(Y|s)]^{-1}I(Y<t,\ \delta_2=1)] + E_{Y|X>s'}[[H(Y|s)H(Y|s')]^{-1}I(Y<t\wedge t',\delta_2=1)]$$

$$= [H(s,0)]^{-1}[IV_a + IV_b + IV_c + IV_d].$$

By integration by parts,

$$IV_a = \int_0^t H(y|s')\ g_{s'}(y\wedge t')[H(y|s)]^{-2}dH_1(y|s) + \int_0^{t'} H(y|s')\cdot g_s(Y\wedge t)[H(y|s')]^{-2}dH_1(y|s'),$$

$$= IV_{a1} + IV_{a2}.$$

It is easy to show that,

$$IV_b = -\int_0^{t'} g_s(Y\wedge t)[H(y|s')]^{-1}dH_1(Y|s') = -IV_{a2},$$

$$IV_c = -\int_0^{t'} g_{s'}(Y\wedge t')[H(y|s)]^{-1}dH_1(y|s'),$$

$$IV_d = -\int_0^{t\wedge t'} [H(y|s)H(y|s')]^{-1}\ dH_1(y|s').$$

Hence $IV = [H(s,0)]^{-1}[IV_{a1} + IV_c + IV_d]$, for $s < s'$.

The proof is now completed. □

9. Appendix III: Proof of Lemma 4.1.

Since $\eta(s,t)$ is uniformly bounded random variables for (s,t) in $[0,S]\times[0,T]$, let M be this uniform upper bound.

Let $\epsilon = H(S,T) > 0$.

$$E([X(E)]^2) \leq 4M\, E|X(E)|$$

$$= 4M\, E|\eta(s,t) - \eta(s,t') - \eta(s',t) + \eta(s',t')|$$

$$= 4M\, E|[H(s,0)]^{-1}I(X>s)[\xi_s(Y,\delta_2,t) - \xi_s(Y,\delta_2,t')] - [H(s',0)]^{-1}I(X>s')\cdot$$

$$[\xi_{s'}(Y,\delta_2,t) - \xi_{s'}(Y,\delta_2,t')]|$$

$$\leq 4M\, E|[H(s,0)]^{-1}I(X>s)[g_s(Y\wedge t) - g_s(Y\wedge t')] - [H(s',0)]^{-1}I(X>s')[g_{s'}(Y\wedge t)$$

$$- g_{s'}(Y\wedge t')]| + 4M\, E|[H(s,0)]^{-1}I(X>s)I(t<Y<t',\delta_2=1)[H(Y|s)]^{-1}$$

$$- [H(s',0)]^{-1}I(X>s')I(t<Y<t',\delta_2=1)\cdot[H(Y|s')]^{-1}|$$

$$= 4M\,(I + II), \text{ where I, II stands for the first \& second expectation respectively.} \tag{9.1}$$

Let $H_1(s,t) = H_1(t|s)\cdot H(s,0) = P(X > s,\ Y > t,\ \delta_2 = 1)$. Let's consider I first.

$$I = E|I(x>s)\cdot\int_{Y\wedge t'}^{Y\wedge t} [H(s,y)]^{-2}dH_1(s,y) - I(X>s')\int_{Y\wedge t'}^{Y\wedge t}[H(s',y)]^{-2}dH_1(s',y)|$$

$$\leq E|\int_{Y\wedge t'}^{Y\wedge t} [H(s,y)]^{-2}dH_1(s,y) - \int_{Y\wedge t'}^{Y\wedge t} [H(s',y)]^{-2}dH_1(s',y)|$$

$$+ E|I(s<X<s')\int_{Y\wedge t'}^{Y\wedge t} [H(s,y)]^{-2}dH_1(s,y)|$$

$$\leq E|\int_{Y\wedge t'}^{Y\wedge t}([H(s,y)]^{-2}-[H(s',y)]^{-2})dH_1(s,y)| + E|\int_{Y\wedge t'}^{Y\wedge t} [H(s',y)]^{-2}d(H_1(s,y) - H_1(s',y))$$

$$+ E|I(s<X<s') \int_{Y\wedge t'}^{Y\wedge t} [H(s,y)]^{-2} dH_1(s,y)|$$

$$= I_a + I_b + I_c. \tag{9.2}$$

$$I_a \leq 2\varepsilon^{-4} E|\int_{Y\wedge t'}^{Y\wedge t} [H(s,y) - H(s',y)]\, dH_1(s,y)|$$

$$\leq 2\varepsilon^{-4} [H(s,0) - H(s',0)] \cdot E|H_1(s,Y\wedge t) - H_1(s, Y\wedge t')|$$

$$\leq 2\varepsilon^{-4} [H(s,0) - H(s',0)] \cdot E[H_1(0,Y\wedge t) - H_1(0, Y\wedge t')|$$

$$\leq 2\varepsilon^{-4} [H(s,0) - H(s',0)]\, [H_1(0,t) - H_1(0,t')] \tag{9.3}$$

$$I_b \leq \varepsilon^{-2} E|H_1(s,Y\wedge t) - H_1(s',Y\wedge t) - H_1(s,Y\wedge t') + H_1(s', Y\wedge t')|$$

$$\leq \varepsilon^{-2} [H_1(s,t) - H_1(s',t) - H_1(s,t') + H_1(s',t')]. \tag{9.4}$$

$$I_c \leq \varepsilon^{-2} E|I(s<X<s') \cdot [H_1(s,Y\wedge t) - H_1(s,Y\wedge t')]|$$

$$\leq \varepsilon^{-2} E|I(s<X<s') \cdot [H_1(0,Y\wedge t) - H_1(0,Y\wedge t')]|$$

$$\leq \varepsilon^{-2} [H(s,0) - H(s',0)][H_1(0,t) - H_1(0,t')] \tag{9.5}$$

Next, consider II

$$II = E|I(t<Y<t',\ \delta_2 = 1)[I(X>s)[H(s,Y)]^{-1} - I(X>s')[H(s',Y)]^{-1}]|$$

$$\leq \varepsilon^{-2} E|I(t<Y<t',\ \delta_2 = 1) \cdot [H(s',Y) - H(s,Y)]| + \varepsilon^{-2} E|I(t<Y<t',\ \delta_2 = 1) I(s<X<s')|$$

$$\leq \varepsilon^{-2}[H(s,0) - H(s',0)]\cdot H_1(0,t) - H_1(0,t')] + \varepsilon^{-2}[H_1(s,t) - H_1(s,t') - H_1(s',t) + H_1(s',t')] \tag{9.6}$$

For any x, y

let $W(x,y) = W((x,\infty)x(y,\infty))$

$$= 4M\,[2\varepsilon^{-4} + 2\varepsilon^{-2}]H(x,0)H_1(0,y) + 8M\varepsilon^{-2}H_1(x,y).$$

It now follows from (9.1) to (9.6) that,

$E([X(E)]^2) \leq W(E)$. □

Noted that we have used the same notation W for both the measure and the survival function.

Acknowledgement: Part of this work was done while the first author visited the Department of Statistics at the University of California, Berkeley. The authors would like to thank Professor Kjell Doksum and Professor Lucien Le Cam for helpful discussions.

References

Bickel, P.J. and Freedman, D. (1981). Some asymptotic theory for the bootstrap. Ann. Statist. 9, 1196-1217.

Bickel, P.J. and Wichura, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. Ann. Math. Statist. 42, 1656-1670.

Billingsley, P. (1968). Convergence of Probability Measures, Wiley.

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. Ann. Statist. 2, 437-443.

Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. Biometrika, 68 (2), 417-423.

Campbell, G. (1982). Asymptotic properties of several nonparametric multivariate distribution function estimators under random censoring. Survival Analysis, Ed. by John Crowley and Richard A. Johnson. IMS Lecture notes, Vol. 2, 243-256.

Campbell, G. and Földes, A. (1982). Large sample properties of nonparametric bivariate estimators with censored data. Proceedings of the International Colloquia on Nonparametric Statistical Inference, Budapest 1980, 23-28, North Holland.

Csörgő, S. and Horváth, L. (1983). The rate of strong uniform consistency for the product-limit estimator. Z. Wahrsch. Verw. Gebiete 62, 411-425.

Dudley, R.M. and Philipp, W. (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. Z. Wahrsch. Verw. Gebiete 62, 509-552.

Efron, B. (1981). Censored data and the bootstrap. J. Amer. Statist. Assoc., 312-321.

Gill, R. (1983). Large sample behavior of the product-limit estimator on the whole line. Ann. Statist., 11, 49-58.

Horváth, L. (1983). The rate of strong uniform consistency for the multivariate product-limit estimator. J. Multivar. Anal. 13, 202-209.

Hanley, J.A. and Parnes, M.N. (1983). Nonparametric estimation of a multivariate distribution in the presence of censoring. Biometrics 39, 129-139.

Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. J. Amer. Statist. Assoc., 53, 457-481.

Kiefer, J. (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm, Pacific J. of Math., 11, 649-660.

Korwar, R.H. and Dahiya, R.C. (1982). Estimation of a bivariate distribution function from incomplete observations. Comm. Statist. - Theor. Meth., 11, (8), 887-897.

Lo, S.H. and Singh, K. (1985). The product-limit estimator and the bootstrap: Some asymptotic representation. Accepted by Z. Wahrsch. Verw. Gebiete. Technical Report, Dept. of Statistics, Rutgers University.

Miller, R.G. (1981). Survival Analysis. Wiley, New York.

Muñoz, A. (1980a). Nonparametric estimation from censored bivariate observations. Technical Report No. 60, Division of Biostatistics, Stanford University.

Muñoz, A. (1980b). Consistency of the self-consistent estimator of the distribution function from censored observations. Technical Report No. 61, Division of Biostatistics, Stanford University.

Singh, K. (1981). On asymptotic accuracy of Efron's bootstrap. Ann. Statist. 9, 1187-1195.

Tsai, W.Y., Leurgans, S. and Crowley, J. (1983). Nonparametric estimation of a bivariate survival function in the presence of censoring. Technical Report 34160, Brookhaven National Laboratory.

END

10-86

DTIC